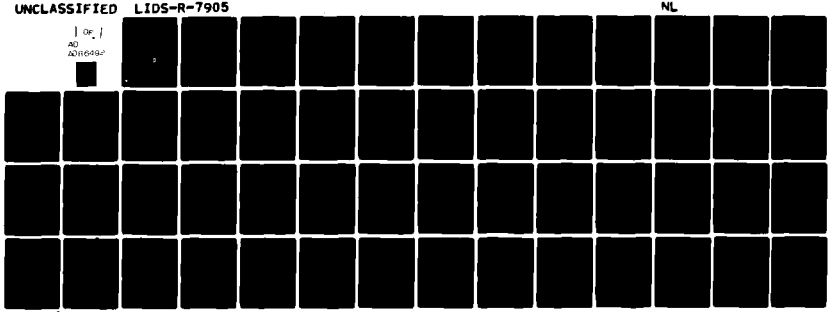
AD-A086 492 MASSACHUSETTS INST OF TECH CAMBRIDGE LAB FOR INFORMA--ETC F/G 12/1
ON-LINE TUNING OF MULTIVARIABLE PI CONTROLLERS USING PRINCIPAL --ETC(U)
APR 79 B C MOORE N00014-79-C-0377
UNCLASSIFIED LIDS-R-7905 NL



END
DATE FILMED
8-80
DTIC

(6)

ON-LINE TUNING OF MULTIVARIABLE PI CONTROLLERS USING PRINCIPAL COMPONENT ANALYSIS: PRELIMINARY RESULTS

by

(10) B.C. Moore

(14) LIDS-P-1045

Systems Control Report No. 7905

(11) April 1979

(12) 54

*Department of Electrical Engineering, University of Toronto, Toronto, Ontario, Canada. M5S 1A4

This research was supported by the Natural Sciences and Engineering Research Council of Canada, under operation grant No. A3925.

ON-LINE TUNING OF MULTIVARIABLE PI CONTROLLERS USING PRINCIPAL COMPONENT ANALYSIS: PRELIMINARY RESULTS

by

B.C. Moore*

ABSTRACT

There has been considerable interest lately in the application of singular value analysis in systems theory. The basic ideas, however, were developed in statistics (Hotelling introduced principal component analysis in 1933) and are currently used in numerical analysis and digital filtering. The fundamental results underlying principal component analysis are presented in this paper, and these results are applied to the problem of tuning multivariable proportional plus integral controllers. Although the tuning method proposed is preliminary, it is designed to avoid possible traps which would prevent "tight" tuning with conventional tuning of individual loops. When applied to non-interacting loops, the method reduces to conventional tuning of the loops simultaneously.

---

*Department of Electrical Engineering, University of Toronto, Toronto, Ontario, Canada. M5S 1A4

This research was supported by the Natural Sciences and Engineering Research Council of Canada, under operating grant No. A3925.

I. INTRODUCTION

Although there has been considerable interest lately in the application of singular value analysis in systems theory, [1]-[9], the basic analysis techniques involved are at least 46 years old. Hotelling [10], [11], introduced principal component analysis in 1933, and the effectivenss of these techniques was enhanced substantially by the development of an algorithm (S.V.D.), [12], for efficient, accurate computation of the important objects. These techniques are currently used in the numerical analysis [13]-[15], and digital filtering [16]-[18]. Dempster [19] gives an excellent geometric treatment of principal component analysis as well as an overview of its history and relationship to least squares approximation.

In this paper the fundamental results underlying these analysis tools are presented (Section II), and a preliminary method (using these tools) for on-line tuning of multivariable proportional-plus-integral controllers is proposed (Sections III, IV). When the method is applied to a set of non-interacting loops, it reduces to a standard classical technique ([20], page 330).

Although the point will not be pursued here, linearity does not play an essential role in principal component analysis. Preliminary ideas on the application of these tools to nonlinear problems are given in [21].

Notation: $\mathbb{R}$, $\mathbb{C}$ will represent the fields of real and complex numbers. For a vector v and matrix M, v^T, M^T represent the transpose, v^*, M^* represent the complex conjugate and v^H, M^H represent the conjugate transpose.

II. FUNDAMENTAL RESULTS

The results given in this section form the foundation for the principal component analysis techniques to be discussed in later sections. The two propositions given below may be viewed as one result stated in two different contexts: the first involving discrete data samples; the second, continuous data. The constant K plays no role other than an aid for discussion.

Discrete Data:

For convenience the samples are organized into a sequence of vectors $y(1), y(2), \ldots, y(N)$ in $\mathbb{C}^n$, with

$$W^2 \triangleq K \sum_{\ell=1}^{N} y(\ell) y^H(\ell). \qquad (K>0)$$

This matrix is positive semidefinite with a set of non-negative eigenvalues

$$\sigma_1^2 \geq \sigma_2^2 \geq \ldots \geq \sigma_n^2 \geq 0$$

and corresponding mutually orthogonal unit eigenvectors $u_1, u_2, \ldots, u_n$.

Proposition 1A:

Let the scalar sequence $y_i(1), y_i(2), \ldots, y_i(N)$ be defined by

$$y_i(\ell) = u_i^H y(\ell)$$

for $1 \leq i \leq n$. Then $y(\ell) = \sum_{i=1}^{n} u_i y_i(\ell)$, where

$$\left(K \sum_{\ell=1}^{N} |y_i(\ell)|^2 \right)^{1/2} = \sigma_i$$

Proof: The first identity follows from the fact that $U=(u_1,u_2,\ldots,u_n)$ is unitary. Furthermore

$$K \sum_{\ell=1}^{N} \|y_i(\ell)\|^2 = K \sum_{\ell=1}^{N} u_i^H y(\ell) y(\ell)^H u_i = u_i^H W^2 u_i$$

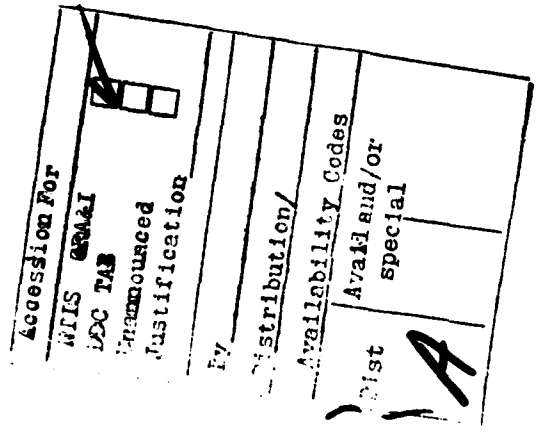
and the second identity follows trivially. □

Stated in other words, Proposition 1A says that we can decompose the vector sequence into the vector sum of spacially orthogonal sequences (called components) ordered with respect to magnitude.

The important objects in Proposition 1A are the pairs (σ_i,u_i), i=1,...,n, and these can be computed without first computing W^2. If a data matrix Y is constructed with N columns consisting of the vectors y(1),y(2),...,y(N), then

$$W^2 = YY^H$$

and (σ_i,u_i), $1 \le i \le n$, are the singular values and left singular vectors of Y. These may be computed directly using the well known algorithm (S.V.D.) developed by Golub and Reinsch [12] (see also [9]).



Continuous Data:

Now consider a piecewise continuous map $f:\mathbb{C}\to C^n$ defined by $y=f(x)$, and let

$$W^2 = K\int_{x_1}^{x_2} f(x)r^H(x)dx \qquad (K>0)$$

with eigenvalues $\sigma_1^2 \geq \sigma_2^2 \geq \ldots \geq \sigma_n^2 \geq 0$ and mutually orthogonal unit eigenvectors $u_1, u_2, \ldots, u_n$.

Proposition 1B:

Let $f_i:\mathbb{C}\to\mathbb{C}$ be defined by $f_i(x)=u_i^H f(x)$. Then

$$f(x) = \sum_{i=1}^{n} u_i f_i(x)$$

where

$$\left[K\int_{x_1}^{x_2} |f_i(x)|^2 dx\right]^{1/2} = \sigma_i$$

Proof: Similar to the proof of Proposition 1A. □

Here we decompose the vector valued function into the sum of spacially orthogonal component functions ordered with respect to magnitude. The same computational tool, S.V.D., can be used to compute (σ_i, u_i) without actually computing W^2. To do this, divide $[x_1, x_2]$ into N evenly spaced sample points and construct a data matrix Y whose columns are the samples. For N large, rectangular approximation of the integral gives

$$\int_{x_1}^{x_2} f(x)f^H(x)dx \approx (1/N)YY^H$$

To compute (σ_i,u_i), $1\leq i\leq n$, one may apply S.V.D. to the scaled data matrix

$$(K/N)^{1/2}Y$$

A Combined Result for Linear Systems

With linear systems, we shall often encounter a matrix $F(x)$ which can be viewed as a set of maps $f_1(x),f_2(x),\ldots,f_m(x)$, one corresponding to each column of F. In this situation it is appropriate to combine Propositions IA, IB. Let

$$W^2 = K \sum_{i=1}^{m} \int_{x_1}^{x_2} f_i(x)f_i^H(x)dx = K \int_{x_1}^{x_2} F(x)F^H(x)dx.$$

If $(\sigma_1^2,u_1),\ldots,(\sigma_n^2,u_n)$ are the ordered eigenvalue, eigenvector (orthonormal) sets for W^2, and

$$F_i(x) \triangleq u_i^H F(x)$$

then

$$F(x) = \sum_{i=1}^{n} u_i^H F_i(x)$$

where

$$\left(K \int_{x_1}^{x_2} \|F_i(x)\|^2 dx\right)^{1/2} = \sigma_i$$

To compute (σ_i, u_i) in this situation, divide $[x_1, x_2]$ into N evenly spaced sample points $s_1, s_2, \ldots, s_N$, and let

$$Y = (F(s_1) \quad F(s_2) \quad \ldots \quad F(s_N)).$$

If there are m columns in F and N sample points, then Y has mN columns.

Remarks about Perturbations:

Suppose the matrix F(x) is perturbed to give $F_\Delta(x)=F(x)+\Delta F(x)$. Then

$$u_i^H F_\Delta(x) = F_i(x) + u_i^H \Delta F(x)$$

and it follows that

$$\left(\int_{x_1}^{x_2} \| F_i(x) - u_i^H F_\Delta(x) \|^2 \, dt \right)^{1/2} \leq \left(\int_{x_1}^{x_2} \| \Delta F(x) \|^2 \, dx \right)^{1/2}$$

where the equality is achieved if $\Delta F(x)$ is aligned with u_i.

This is a double edged sword. First, it gives a tool for coping with structural instability associated with many theoretical results. Theory says, for example, that if every column of F(x) is contained in a proper subspace S for all $x \in [x_1, x_2]$, then we may project onto S to simplify the situation (by reducing dimensionality). Such a subspace S may, however, be structurally unstable: there may exist arbitrarily small admissable perturbations such that the columns of $F(x)+\Delta F(x)$ are not contained in a proper subspace. We are guaranteed, however, that $F_\Delta(x)$ will have weak components (assuming small perturbations) and that the strong components will define a subspace which is close to S.

The other edge of the sword gives us help in deciding roughly how accurate we can expect the components to be. If for example, there are components whose magnitudes are of the same order as the instrumentation precision, then they may be in error by ≈100% and one can hardly use them (say for feedback) with confidence.

General Comments about Applications:

The results of this section provide a strong tool for spacial analysis (possibly on-line in many situations) of multivariable signals in the time or frequency domains.

If, for example, f(t) is a vector of signals and

$$W^2 = K \int_{t_1}^{t_2} f(t) f^T(t) dt$$

then Proposition 1B gives a decomposition into components ordered with respect to their energy (K=1) or average power $\left(K = \frac{1}{t_2 - t_1}\right)$

$$\left(K \int_{t_1}^{t_2} f_i^2(t) dt \right)^{1/2} = \sigma_i$$

For a vector f(jw) in the frequency domain, and

$$W^2 = \int_{w_1}^{w_2} f(jw) f^H(jw) dw,$$

one gets information about the spacial distribution of f(jw) over the frequency band $[w_1, w_2]$.

Remarks about Computation:

Although one might be alarmed at the thought of using a minicomputer to compute the singular values and left singular vectors of a matrix Y with, say, 10 rows and 100 columns, it is quite reasonable. One can recursively (treating one column of data at a time) reduce Y to a unitarily equivalent matrix (see [14], p. 383)

$$\hat{Y} = YQ \quad (Q \text{ unitary})$$

where $\hat{Y}$ is 10×10; this process requires $\approx 100(10^2)=10{,}000$ operations. Singular value decomposition of $\hat{Y}$ requires $\approx 6(10^3)=6000$ operations, giving a total of operation count of $\approx 16{,}000$.

At first it may seem simpler to find the components by computing the eigenvalues and eigenvectors of YY^H. Experts recommend that this be avoided (by using S.V.D.) for the following reason (see [14], p. 382 for more discussion): This method doubles the demand for internal computer resolution associated with the algorithm.

Specifically, suppose one has invested money in 12 bit A-D converters and has interfaced these properly so that there is 12 bit resolution associated with the data. To get this same resolution on the span $[0,\sigma_1]$; that is, to have

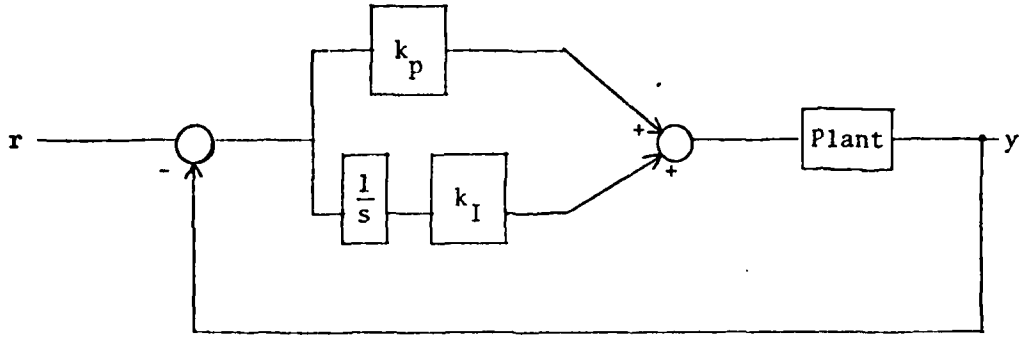
$$\sigma_i - 2^{-12}\sigma_1 \leq \{\text{computed value of } \sigma_i\} \leq \sigma_i + 2^{-12}\sigma_1$$

requires at least 12 bit internal resolution using S.V.D., and at least 24 bit internal resolution using the "squared up" version where YY^H is computed.

Stated in other words, possibly more to the point with current minicomputer hardware organization, it is highly probable (work needs to be done here) that one can get 12 bit resolution with 12 bit A-D converters and 16 bit operations if S.V.D. is used. It is impossible to do so with the other algorithm; one would probably need to carry out the computations using 32 bit arithmetic.

III. MULTIVARIABLE PI CONTROL

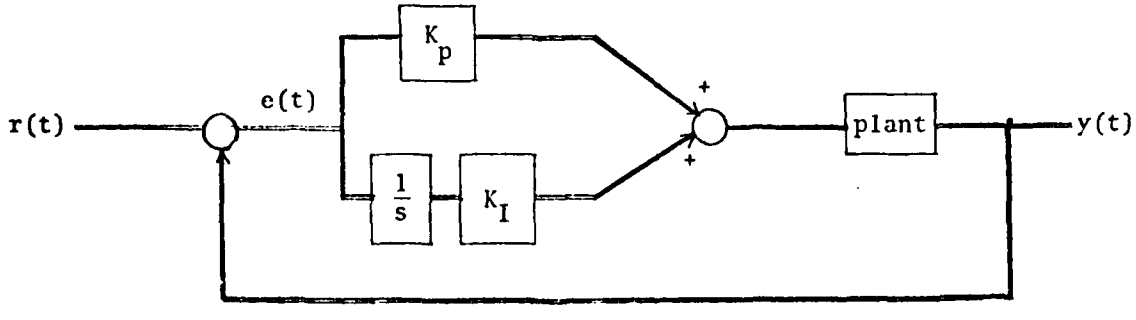
Consider the following PI (Proportional+Integral) control loop which is assumed to be open loop stable:



With a single isolated loop such as this, one can often follow a simple procedure to adjust the gains k_p, k_I; i.e. to "tune" the control loop. One classical method (see [20], p. 330) is the following:

- with k_I=0 increase k_p (this improves response speed) until the step response is highly oscillatory. Reduce k_p by a factor of 2.
- increase k_I (this reduces offset) until the step response is highly oscillatory. Reduce k_I by a factor of 2.

With multiple interacting PI control loops, there are inherent "traps" associated with extending simple procedures such as this. To bring some of these problems into closer view, consider the system shown below



where K_p, K_I are diagonal matrices of proportional and integral gains, respectively. Let's assume that each variable is scaled so that one unit corresponds to a fixed percentage of "full swing" and that we are free to sample e(t), the error vector.

Let E(t) be the matrix made up of error responses to unit steps; i.e. the i^{th} column of E(t) is the error response to a unit step applied to the i^{th} reference input. For a linear system, the response to an arbitrary vector of steps

$$r(t) = r\,\delta(t)$$

is given by

$$e(t) = E(t)r.$$

The steady state error (assuming stability) is

$$e_{ss} = E_{ss}r$$

where $E_{ss} \triangleq \lim_{t\to\infty} E(t)$.

A major trap follows from the fact that an operator sees only <u>projections</u> of the vector e(t) on the basis vectors of a <u>fixed</u> coordinate system associated with the physical arrangement of hardware (sensors). It is possible for rather simple mechanisms to appear complicated in this fixed coordinate system. The following paragraphs show that simple mechanisms involving the notions of settling time, oscillations, and steady state errors may be confusing when viewed through projections.

Settling Time:

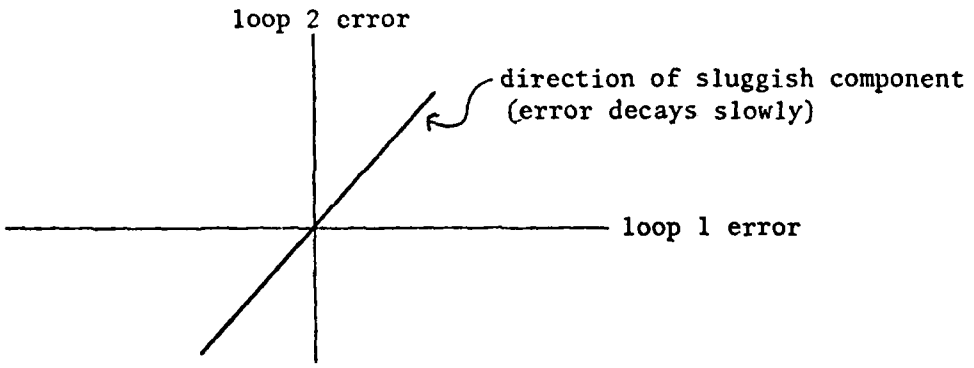
Let $T(t) = E(t) - E_{ss}$ be the transient response map so that

$$e(t) = e_{tr}(t) + e_{ss}$$

where

$$e_{tr}(t) = T(t)r$$

It is certainly possible for the response to have one sluggish component which projects significantly onto each coordinate; i.e. all loops appear sluggish.



Actually for a system with n loops, one can define n settling times as follows: Let

$$W^2(t) = \int_t^\infty T(\tau)T^T(\tau)d\tau$$

with eigenvalue, eigenvector pairs $(\sigma_1^2(t),u_1(t)),(\sigma_2^2(t),u_2(t)),\ldots,(\sigma_n^2(t),u_n(t))$. Then

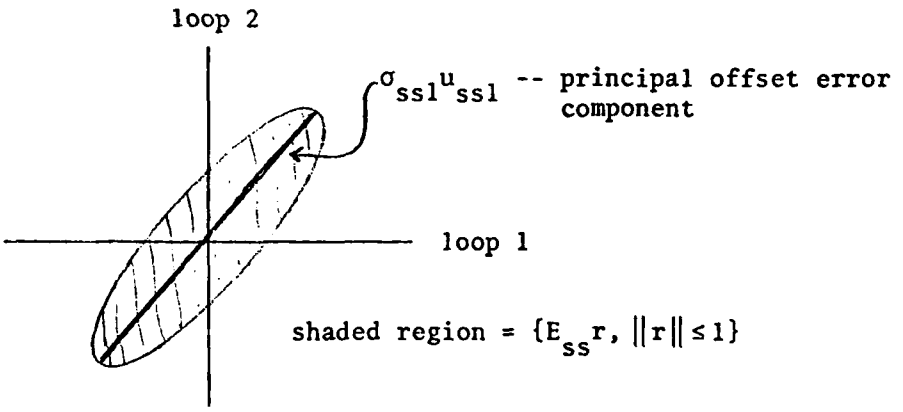
$t_{si} \triangleq$ minimum time such that $\sigma_i(t_{si}) < \varepsilon$.

Note that $t_{s1} \geq t_{s2} \geq \ldots \geq t_{sn}$.

Further on in this section we shall propose one procedure for "tightening" the response in the direction associated with sluggish components.

Steady State Errors:

Consider $W_{ss}^2 \triangleq E_{ss}E_{ss}^T$ with eigenvalue, eigenvector pairs $(\sigma_{ss1}^2, u_{ss1}), \ldots, (\sigma_{ssn}^2, u_{ssn})$. It is entirely possible that there are only a few strong offset error components which project onto every loop



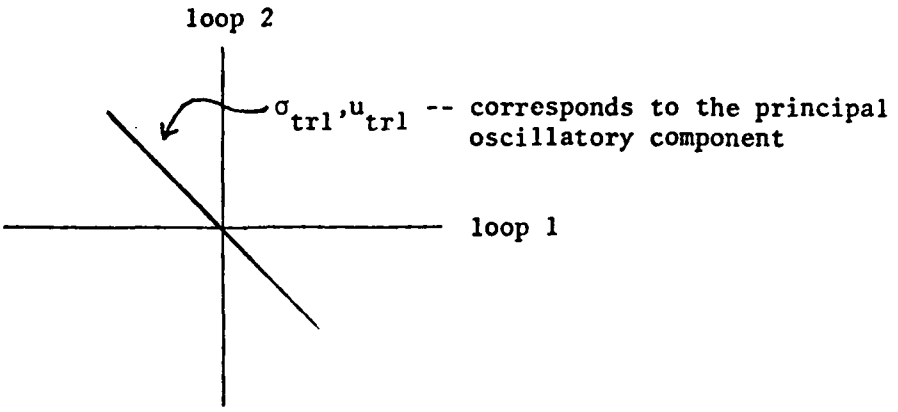
Oscillations:

To avoid unnecessary complications, let's assume that there is a single lightly damped sinusoidal component observed in the loop error responses, and that the observed frequency is w. Let $T_{av} = \int_0^{2\pi/w} T(t)dt$ and

$$W_{tr}^2 = \int_0^{2\pi/w} (T(t)-T_{av})(T(t)-T_{av})^T dt$$

with eigenvalue, eigenvector pairs $(\sigma_{tr1}^2, u_{tr1}), \ldots, (\sigma_{trn}^2, u_{trn})$.

Again it is possible that there is one strong oscillatory component which projects onto a number of loops



A tuning method which allows one to avoid these traps is given in the next section.

IV. A PRELIMINARY ON-LINE TUNING METHOD

These ideas are preliminary and essentially untested. Undoubtedly tests currently in progress will lead to better tuning methods than the rather vague one proposed in the following paragraphs. The tuning method is basically the classical technique (given in Section III) applied to principal components. If loops are noninteracting, the method reduces to the classical technique applied to all loops simultaneously.

Part 1 Tightening the Response:

Let's assume that initially $K_I=0$, and that the diagonal components of K_p are increased (at the same rate, if you like) until there is a highly oscillatory component. If the components of

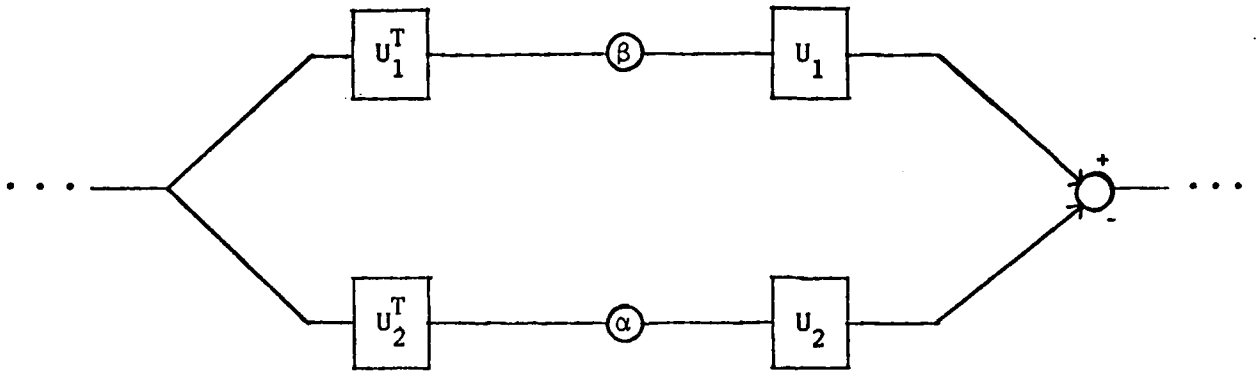
$$T(t) - T_{av}$$

have magnitudes $\sigma_{trl},\ldots,\sigma_{trn}$ which are roughly equal, then this step is essentially complete; reduce the gains by a factor of two (to reduce the oscillations) and proceed to Part 2.

If there are some weak components, then increase gains in these directions. Specifically, suppose $\sigma_{tk} >> \sigma_{trk+1}$, and let

$$U_1 = (u_1 \ u_2 \ \ldots \ u_k); \ U_2 = (u_{k+1} \ \ldots \ u_n).$$

Then in the proportional branch, we insert



where α,β are tuning parameters with $\alpha=\beta=1$ initially. The parameter α should be increased until oscillations appear in the lower branch; it may be necessary to reduce β in the process to maintain stability; i.e. to keep upper branch oscillations bounded.

This process

- compute components of $T(t)-T_{av}$
- insert a Rotation-Tuning Gains-Rotation block
- adjust gains to get oscillations in weak branch

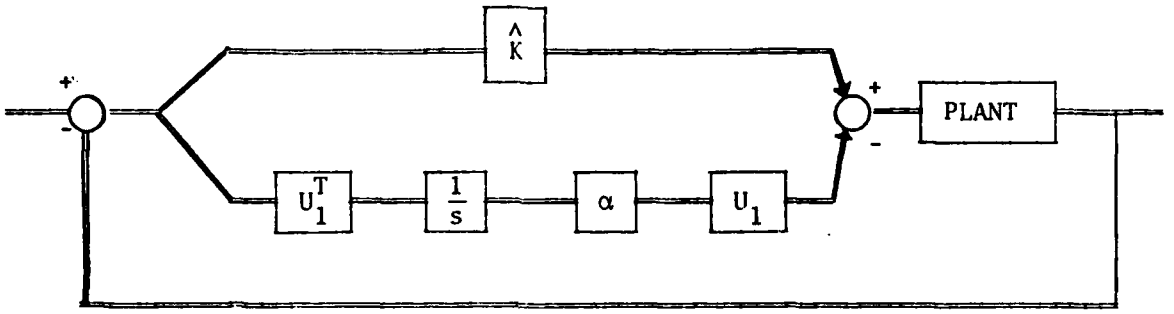
can be continued until the oscillations are "full" with no weak components. The resulting gain matrix $\hat{K}_p$ in the proportional branch is the product of K_p and the inserted blocks. The elements of $\hat{K}_p$ should be reduced by a factor of 2 to reduce oscillations. This completes Part 1.

Part 2 -- Inserting integral gain to reduce offset errors:

At this point it may be necessary to insert integral action if there is significant offset. Here we shall use the components of E_{ss} (with controller $\hat{K}$). If $\sigma_{ss\,k} >> \sigma_{ss\,k+1}$, let

$$U_1 = (u_{ss1} \quad u_{ss2} \quad \cdots \quad u_{ssk}); \quad U_2 = (u_{ssk+1} \cdots u_{ssn})$$

and configure the controller in the following way:

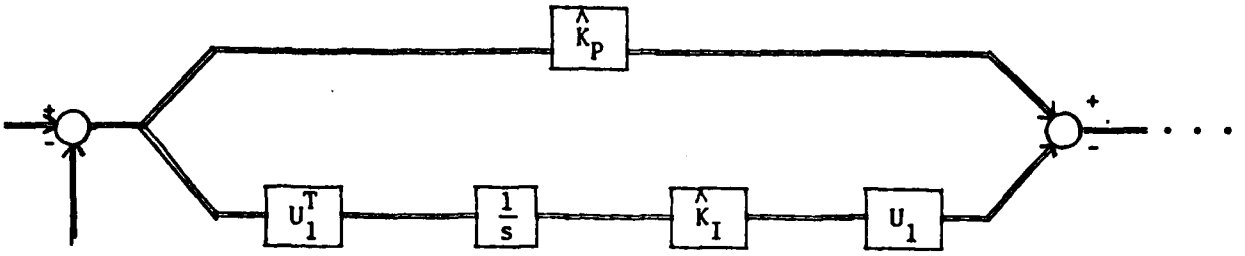


with this configuration, α should be increased from zero until oscillations occur. If the oscillations are "full" in the space corresponding to the integral branch, reduce α by a factor of two and stop.

If the oscillations in the integral branch are not full, we may proceed as follows iteratively until the oscillations are full.

- compute components of $U_1^T(T(t)-T_{av})U_1$
- insert a Rotation-Tuning Gains-Rotation block (as in part 1) in the integral branch
- adjust gains to increase oscillation in weak integral branch

We are one step away from completion of the proposed tuning algorithm, with the following controller structure



The last step is to reduce the elements of $\hat{K}_I$ by a factor of two.

Concluding Remarks:

It is certainly possible in many process control problems to use a single microprocessor system (<$10,000) for multivariable PI control of a small number of loops (say 10). Process control systems of this type have been applied successfully (see [22], [23]) in the control of industrial heating and air conditioning equipment (boilers, chillers, cooling towers, air handlers, etc.). The tuning method proposed in this section can be implemented with little or no additional hardware. The procedure is simple and consistent with classical tuning methods. With a well designed man-machine interface (say bar graphs for singular values and automatic generation of rotation blocks) the method would probably be acceptable to plant operators.

[1] Moore, B.C. and A.J. Laub, "Computation of supremal (A,B)-invariant and controllability subspaces", IEEE Trans. on Auto. Control, vol. AC-23, no. 5, pp. 783-792.

[2] Moore, B.C., "Singular value analysis of linear systems", Proceedings of the 1978 IEEE Conference on Decision and Control, Jan. 1979, pp. 66-73. Revised paper to appear in IEEE Trans. on Auto. Control, as "Principal Component Analysis in Linear Systems".

[3] Doyle, J.C., "Robustness of multiloop linear feedback systems", Proceedings of the 1978 IEEE Conference on Decision and Control, Jan. 1979, pp. 12-17.

[4] Stein, G. and J. Doyle, "Singular values and feedback: design examples", Proceedings of the 16th Annual Allerton Conference on Communication, Control, and Computing, Oct. 1978, pp. 461-470.

[5] Sandell, N.R., "Robust stability of multivariable feedback systems", Proceedings of the 16th Annual Allerton Conference on Communication, Control, and Computing, Oct. 1978, pp. 471-479.

[6] Safonov, M.G., "Tight bounds on the response of multivariable systems with component uncertainty", Proceedings of the 16th Annual Allerton Conference on Communication, Control, and Computing, Oct. 1978, pp. 451-459.

[7] Laub, A.J., "Computational aspects of the singular value decomposition and some applications", Proceedings of the 16th Annual Allerton Conference on Communication, Control, and Computing, Oct. 1978, pp. 432-442.

[8] Laub, A.J., "Linear multivariable control: numerical considerations", Electronic Systems Laboratory (Now Laboratory for Information and Decision Systems) report ESL-P-833, Massachusetts Institute for Technology, Cambridge, Mass. 02139.

[9] Klema, V.C. and A.J. Laub, "The singular value decomposition: its computation and some applications", Laboratory for Information and Decision Systems, report LIDS-R-829, Massachusetts Institute of Technology, Cambridge, Mass. 02139

[10] Hotelling, H., "Analysis of a complex of statistical variables into principal components", Journal of Educational Psychology, vol. 24 (1933), pp. 417-441, pp. 498-520.

[11] Hotelling, H., "Simplified calculation of principal components", Psychometrika, vol. 1, (1936), pp. 27-35.

[12] Golub, G.H. and C. Reinsch, "Singular value decomposition and least squares solutions", Numer. Math., 14 (197), pp. 403-420.

[13] Forsythe, G.E. and C.B. Moler, Computer Solution of Linear Algebraic Systems, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1967.

[14] Stewart, G.W., Introduction to Matrix Computations, Academic Press, New York, 1973.

[15] Lawson, C.L. and R.J. Hanson, Solving Lieast Squares Problems, Prentice-Hall, Englewood Cliffs, N.J., 1974.

[16] Andrews, H.C. and B.R. Hunt, Digital Image Restoration, Prentice-Hall, 1977.

[17] Mullis, C.T. and R.A. Roberts, "Synthesis of minimum roundoff noise fixed point digital filters", IEEE Trans. on Circuits and Systems, vol. CAS-23, no. 9, pp. 551-562.

[18] Mills, W.L., C.T. Mullis and R.A. Roberts, "Digital filter realizations without overflow oscillations", IEEE Trans. on Acoustics, Speech, and Signal Processing, vol. ASSP-26, no. 4, pp. 334-338.

[19] Dempster, A.P., Elements of Continuous Multivariate Analysis, Addison-Wesley, 1969.

[20] Luyben, W.L., Process Modelling, Simulation, and Control for Chemical Engineers, McGraw-Hill, 1973.

[21] Moore, B.C., "Principal component analysis in nonlinear systems: preliminary results", University of Toronto Systems Control Report No. 7903, April 1979, Department of Electrical Engineering, University of Toronto, Toronto, Ontario, Canada. M5S 1A4.

[22] Meals, M. and C.F. Moore, "Microprocessor based distributed control — software organization", presented at the 1978 Conference on Computers, Electronics, and Control, Nov. 1978, Toronto, Ontario, Canada. M5S 1A4.

[23] Jeffreys, J., R. Gaddis, and C.F. Moore, "Microprocessor based distributed control — hardware organization", presented at the 1978 Conference on Computers, Electronics, and Control, Nov. 1978, Toronto, Ontario, Canada. M5S 1A4.

ROBUST STABILITY OF LINEAR SYSTEMS - SOME COMPUTATIONAL CONSIDERATIONS*

by

Alan J. Laub**

1. INTRODUCTION

In this paper we shall concentrate on some of the computational issues which arise in studying the robust stability of linear systems. Insofar as possible, we shall use notation consistent with Stein's paper [1] and we shall make frequent reference to that work.

As we saw in [1] a basic stability question for a linear time-invariant system with transfer matrix G(s) is the following: given that a nominal closed-loop feedback system is stable, does the feedback system remain stable when subjected to perturbations and how large can those perturbations be? It turned out, through invocation of the Nyquist Criterion, that the size of the allowable perturbations was related to the "nearness to singularity" of the return difference matrix $I + G(j\omega)$. Closed-loop stability was said to be "robust" if G could tolerate considerable perturbation before I + G became singular.

---

* Invited Paper presented at the Second Annual Workshop on the Information Linkage between Applied Mathematics and Industry held at Naval Postgraduate School, Monterey, California, Feb. 20-24, 1979; this research was partially supported by NASA under grant NGL-22-009-124 and the Department of Energy under grant ET-78-(01-3395).

** Laboratory for Information and Decision Systems, Rm. 35-331, M.I.T., Cambridge, MA 02139.

We shall now indulge in a modicum of abstraction and attempt to formalize the notion of robustness. The definition will employ some jargon from algebraic geometry and will be applicable to a variety of situations. While no deep results from algebraic geometry need be employed, the exercise of formulating a precise definition is a useful one for clarifying one's thinking.

Let $p \in \mathbb{R}^N$ be a vector of parameters from some problem being studied and suppose we are interested in some property Π of this data. The vector p may consist of the elements of various matrices, for example. If Π is true at some nominal parameter set p_0 we are frequently concerned with whether Π remains true in a "neighborhood" of p_0.

For example, p_0 may be the elements $(a_{11}, \ldots, a_{1n}, a_{21}, \ldots, a_{nn})$ of a nonsingular nxn matrix A_0 and we are interested in the nonsingularity of nearby matrices. We shall proceed to formalize the often-heard statement that "almost all nxn matrices are nonsingular". First, the jargon:

Definition 1: A variety $V = \{p \in \mathbb{R}^N: \psi_i(p_1, \ldots, p_N) = 0, i = 1, \ldots, k\}$ where $\psi_i(x_1, \ldots, x_N) \in \mathbb{R}[x_1, \ldots, x_N]$ are polynomials. V is proper if $V \neq \mathbb{R}^N$ and nontrivial if $V \neq \phi$.

Definition 2: A property is a function $\Pi: \mathbb{R}^N \to \{0, 1\}$. The property Π holds if $\Pi(p) = 1$ and fails if $\Pi(p) = 0$.

Definition 3: If V is a proper variety, Π is generic relative to V provided $\Pi(p) = 0$ only if $p \in V$. A property Π is generic if such a V exists.

Our discussion to this point is purely algebraic. Now let us introduce a topology on $\mathbb{R}^N$, say the topology induced by some vector norm $\|\cdot\|$.

Furthermore, let V be any nontrivial, proper variety. Then we have the following topological definition.

Definition 4: The property Π is well-posed at $p \in V^c$ (the complement of V) if Π also holds in a sufficiently small neighborhood of p.

Lemma 1: The set S of points where a generic property is well-posed is open and dense. Moreover, the Lebesgue measure of S^c is zero.

The proof of Lemma 1 is routine and is omitted. It is easy to see that a point p where a generic property holds is well-posed but that the converse is not necessarily true.

We now have sufficient framework to make a formal definition of robustness.

Definition 5: Given a point p with generic property Π (generic with respect to some proper variety V) well-posed at p, let

$$d = \min_{v \in V} \| p - v \| .$$

We say Π is robust at p if d is "large".

The number d is frequently difficult to compute or estimate. When it can be determined, it gives valuable information about how much perturbation or uncertainty can be tolerated at p. For the situation of special interest in this paper, Example 2 below, we shall see that d can be explicitly calculated, at least theoretically. We now illustrate the above concepts with two examples.

Example 1

This example is chosen from Wonham [2] who uses the concepts of genericity and well-posedness in nontrivial ways for a variety of control-theoretic problems. In this trivial example, we seek solutions of the system of linear equations

$$Ax = b$$

where $A \in \mathbb{R}^{m \times n}$ (i.e., A is an mxn matrix with real coefficients) and $b \in \mathbb{R}^m$. Our parameter vector is p where

$$p^T = (a_{11}, \ldots, a_{1n}, \ldots, a_{mn}; b_1, \ldots, b_m) \in \mathbb{R}^N, \quad N = mn + m$$

(^{T} denotes transpose). Π is the property of the equation having a solution which is equivalent, of course, to the statements that $b \in \text{Im } A$ or $\text{rk}[A, b] = \text{rk } A$. For example, if $A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$ and $b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$ then

$$\Pi(1,2,2,4;\ b_1,b_2) = \begin{cases} 0 & \text{if } b_2 \neq 2b_1 \\ 1 & \text{if } b_2 = 2b_1 \end{cases}$$

It is then easy to show the following: (see [2])

1. Π is generic if and only if $m \leq n$.
2. Π is well-posed at p if and only if $\text{rk } A = m$.

Example 2

This example is similar to Example 1 in the special case m = n. We are given a nonsingular matrix $A \in \mathbb{R}^{n \times n}$ and we are concerned with the nearness of A to singularity. Identifying A with $p^T = (a_{11}, \ldots, a_{1n}, a_{21}, \ldots, a_{nn})$ we define the property Π by

$$\Pi(p) = \begin{cases} 0 & \text{if p represents a singular matrix} \\ 1 & \text{if p represents a nonsingular matrix .} \end{cases}$$

Then it is easy to see that Π is a generic property and well-posed where it holds. This is the precise statement that "almost all nxn matrices are nonsingular". Formally writing down the determinant of A as a polynomial in $a_{11}, \ldots, a_{nn}$ defines the necessary variety V. It turns out, in a theorem attributed by Kahan [3] to Gastinel, that the distance d from a point $p \in V^c$ to V can be explicitly determined.

Theorem 1: A nonsingular matrix A differs from a singular matrix by no more in norm than $\frac{1}{\|A^{-1}\|}$, i.e., given A,

$$\frac{1}{\|A^{-1}\|} = \min\{\|E\| : A + E \text{ is singular}\} .$$

Thus $d = \frac{1}{\|A^{-1}\|}$ and we might say that A is robust with respect to invertibility if d is "large". To avoid certain scaling difficulties, it may be more desirable to work with a relative measure of distance, d^{rel}, defined by

$$d^{rel} = \frac{d}{\|A\|} = \frac{1}{\|A\| \cdot \|A^{-1}\|} = \frac{1}{\kappa(A)} .$$

The quantity $\kappa(A)$ is recognizable as the condition number of A with respect to inversion. Of course, all the above quantities depend on the particular matrix norm used. To exhibit the specific dependence on the norm $\|\cdot\|_q$ we shall append a subscript "q". For example,

$$d_q = \frac{1}{\|A^{-1}\|_q} \quad .$$

The minimizing E in Theorem 1 can be explicitly constructed for a number of standard matrix norms. For example:

1. $\|A\|_2 = (\lambda_{max}(A^TA))^{1/2}$.

Let A have singular value decomposition $A = USV^T$ where $U, V \in \mathbb{R}^{nxn}$ are orthogonal and $S = \text{diag}\{\sigma_1, \ldots, \sigma_n\}$. The σ_i's, $\sigma_1 \geq \ldots \geq \sigma_n > 0$, are the singular values of A. The minimizing E is given by $E = URV^T$ where $R = \text{diag}\{0, \ldots, 0, -\sigma_n\}$. Then

$$\|E\|_2 = \sigma_n = \frac{1}{\|A^{-1}\|}$$

and A + E is singular. The singular direction, i.e., a nonzero vector z such that $(A + E)z = 0$, is given by the $n^{\underline{th}}$ column of V.

2. $\|A\|_\infty = \max_{i \in \underline{n}} \{\sum_{j=1}^{n} |a_{ij}|\}$, $\underline{n} = \{1,2,\ldots, n\}$.

Suppose $A^{-1} = [\alpha_{ij}]$ and $\|A^{-1}\| = \sum_{j=1}^{n} |\alpha_{kj}|$ for $k \in \underline{n}$. Then the minimizing E is a matrix all of whose elements are 0 except for the $k^{\underline{th}}$ column which consists of the elements

$$\frac{-\text{sgn } \alpha_{k1}}{\|A^{-1}\|_\infty}, \ldots, \frac{-\text{sgn } \alpha_{kn}}{\|A^{-1}\|_\infty} \quad .$$

In fact, letting $z = \text{sgn}\begin{pmatrix} \alpha_{k1} \\ \vdots \\ \alpha_{kn} \end{pmatrix}$ and $u = \begin{pmatrix} 0 \\ \vdots \\ 1/\|A^{-1}\|_\infty \\ \vdots \\ 0 \end{pmatrix}$ with the only

nonzero component of u being in the $k^{\underline{th}}$ row, we have $E = -zu^T$ and clearly $\|E\|_\infty = \frac{1}{\|A^{-1}\|_\infty}$. Now, $(I + EA^{-1})z = (1 - u^TA^{-1}z)z = 0$ since the $k^{\underline{th}}$ element of $A^{-1}z$ is $\sum_{j=1}^{k} |\alpha_{kj}| = \|A^{-1}\|_\infty$ so that $u^TA^{-1}z = 1$. Hence $A + E = (I + EA^{-1})A$ is singular. Moreover, the singular direction is given by $A^{-1}z$ since $(A+E)A^{-1}z = 0$.

3. $\|A\|_1 = \max_{j \in \underline{n}} \{\sum_{i=1}^{n} |a_{ij}|\}$.

The results for this norm are analogous to $\|\cdot\|_\infty$ and can be derived directly or by noticing that $\|A\|_1 = \|A^T\|_\infty$. For completeness we note that if $\|A\|_1 = \sum_{i=1}^{n} |\alpha_{ik}|$ for $k \in \underline{n}$ and

$$z = \operatorname{sgn}\begin{pmatrix} \alpha_{1k} \\ \cdot \\ \cdot \\ \cdot \\ \alpha_{nk} \end{pmatrix}, \qquad u = \begin{pmatrix} 0 \\ \vdots \\ 1/\|A^{-1}\|_1 \\ \vdots \\ 0 \end{pmatrix}$$

then the minimizing E is given by $E = -uz^T$.

We shall see in Section 3 how the results in Example 2 can be applied in studying robustness of stability of linear systems.

2. THE LINEAR SYSTEMS SETTING

In this section we shall provide a brief introduction to both the linear time-invariant systems setting and to the fundamental notion of feedback. This will serve a two-fold purpose: first, to set the stage for the basic stability results and second, to introduce the jargon and notation, especially for non-engineers. This material is standard and can be found in any of a number of standard textbooks on control systems.

We shall consider modelling physical systems by models which take the form of a system of linear constant-coefficient ordinary differential equations

$$\dot{x}(t) = Ax(t) + Bu(t) \tag{1}$$

$$y(t) = Cx(t) \tag{2}$$

The vector x is an n-vector of states, u is an m-vector of inputs or controls, and y is an r-vector of outputs or observed variables.

Starting from the initial condition x(0) the solution of (1) is well-known to be

$$x(t) = e^{tA}x(0) + \int_0^t e^{(t-\tau)A}Bu(\tau)d\tau, \quad t \geq 0 \tag{3}$$

so that the output is given by

$$y(t) = Ce^{tA}x(0) + \int_0^t Ce^{(t-\tau)A}Bu(\tau)d\tau, \quad t \geq 0 \tag{4}$$

where e^{tA} is the matrix exponential defined, but not generally computed, by

$$e^{tA} := \sum_{k=0}^{+\infty} \frac{t^k A^k}{k!} \quad .$$

The matrix $Ce^{tA}B$ is called the impulse response matrix.

Denoting (one-sided) Laplace transforms by upper case letters, take Laplace transforms in (4) to get

$$Y(s) = CX(s) = C(sI - A)^{-1}x(0) + C(sI - A)^{-1}BU(s) . \qquad (5)$$

The matrix $G(s) := C(sI - A)^{-1}B$ is called the transfer matrix. Notice that G(s) is the Laplace transform of the impulse response matrix.

As will be seen in the sequel, it is of interest to study the response of the above linear system to sinusoidal inputs of the form

$$u(t) = e^{j\omega t}v, \quad t \geq 0 \qquad (6)$$

where v is a constant m-vector, ω is the frequency of the sinusoidal input, and $j = \sqrt{-1}$. The response of (1) to this input can then be shown to be of the form

$$x(t) = e^{tA}a + (j\omega I - A)^{-1}Bve^{j\omega t}, \qquad t \geq 0 \qquad (7)$$

where a is a constant n-vector depending on initial conditions. Now, in the case where A is stable (i.e., its spectrum lies in the left-half of the complex plane) the quantity $e^{tA}a$ goes to zero as t approaches $+\infty$. The resulting output

$$y(t) = C(j\omega I - A)^{-1}Bve^{j\omega t} \qquad (8)$$

is called the steady-state frequency response and the matrix

$$G(j\omega) := C(j\omega I - A)^{-1}B , \qquad (9)$$

which turns out to be the transfer function evaluated at $s = j\omega$, is called the frequency response matrix.

Turning now to the case of a real signal given by

$$u_k(t) = v_k \sin(\omega t + \phi_k), \quad t \geq 0 \tag{10}$$

$$u_i(t) = 0, \qquad i = 1, \ldots, m;\ i \neq k,$$

we have steady-state frequency response of the ℓth output given by

$$y_\ell(t) = |G_{\ell k}(j\omega)| v_k \sin(\omega t + \phi_k + \psi_{\ell k}) \tag{11}$$

where $\psi_{\ell k} = \arg(G_{\ell k}(j\omega))$.

Aside from its obvious importance in the above analysis, the frequency response matrix is important for two reasons:

1. Sinusoidal signals are readily available as test signals for a linear system so $G(j\omega)$ can be experimentally determined.

2. Various plots or graphs associated with $G(j\omega)$ can be used to analyze control systems, for example, with respect to stability. Plots such as those associated with the names of Bode, Nichols, and Nyquist are essentially different ways of graphically representing $|G_{\ell k}(j\omega)|$ and $\arg(G_{\ell k}(j\omega))$ as functions of ω. These plots are used extensively in the analysis of single-input single-output control systems where the robustness of stability, e.g., the amount of gain and phase margin available, is checked essentially visually. The appropriate techniques in the multiple-input multiple-output case are still being investigated and part of the motivation for the research in [1] and this paper is directed towards this end.

Turning now to the notion of feedback whose essential idea is to allow for stability of a system in the face of uncertainty (noise, model error, etc.), the diagram below illustrates the basic (unity) feedback control system:

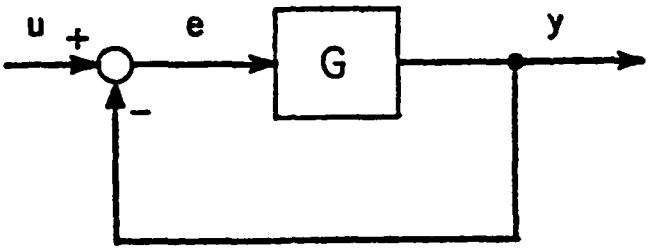


Fig. 1. Basic Feedback Control System

Here u is a reference input, y is the output, and e = u - y is the error or difference between the reference input and the output which we wish to be,ideally, zero. The plant, compensators, actuators, and sensors are all represented by G. There are much more elaborate and detailed feedback structures than that described above and the structure can be studied in a considerably more general function-space setting (see [4], for example) than the simple linear causal time-invariant setting we shall consider. However, the simple system is adequate to exhibit most of the key ideas in this paper. Now, in this system we have

$$e = u - y = u - Ge \tag{12}$$

or,

$$(I + G)e = u \tag{13}$$

The quantity $I + G$ is called the return difference matrix. As in [1], the matrix $G(j\omega)$ then provides sufficient data, via the Nyquist criterion, to test for stability of the closed-loop system. Henceforth, we shall assume that our nominal feedback system above is stable in which case $I+G$ is invertible. Then from (13) we have

$$e = (I+G)^{-1}u \tag{14}$$

so that

$$y = Ge = G(I+G)^{-1}u \ . \tag{15}$$

In (15), the quantity $G(s)(I + G(s))^{-1}$ is called the closed-loop transfer matrix while $G(j\omega)(I+G(j\omega))^{-1}$ is called the closed-loop frequency response matrix. We then pose the basic stability question:

> Does the nominal feedback system remain stable when subjected to perturbations and how large can those perturbations be?

Let us observe at this point that there is nothing sacred about linearity in the above discussion and more general nonlinear treatments can be found in [4] and [5], for example. The question of "nearness to singularity" of $(I+G)$, even in the nonlinear case, is naturally intimately related to a notion of condition number for nonlinear equations. The interested reader could readily adapt the ideas of Rheinboldt [6] to the particular application at hand here.

3. BASIC STABILITY RESULTS AND RELATED TOPICS

a. ADDITIVE AND MULTIPLICATIVE PERTURBATIONS

We shall consider two fundamental types of perturbations in the basic feedback system of Fig. 1. Throughout this section, $\|\cdot\|$ will denote any matrix norm with $\|I\| = 1$. The first case to be considered is the case of additive perturbations to G, pictured below:

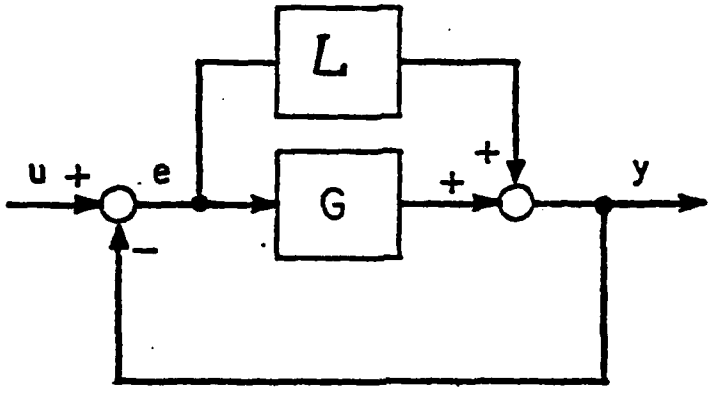


Fig. 2. Additive Perturbations

In other words, the nominal G is perturbed to G + L. Under the assumptions that both the nominal closed-loop system and the perturbation L are stable it can be seen from the Nyquist criterion and the identity

$$I + G + L \equiv (I+G)\ [I + (I+G)^{-1}L] \tag{16}$$

that the perturbed closed-loop system remains stable if

$$\|(I + G(j\omega))^{-1}L(j\omega)\| < 1, \quad \omega > 0 \tag{17}$$

A weaker condition than (17) but one which directly exposes L is

$$\|L(j\omega)\| < \frac{1}{\|(I + G(j\omega))^{-1}\|}, \quad \omega > 0 \tag{18}$$

The second case to be considered is that of multiplicative perturbations:

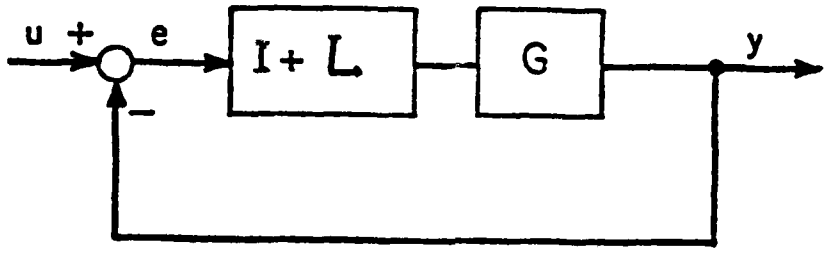


Fig. 3. Multiplicative Perturbations

In this case, the nominal G is perturbed to $G(I + L)$. Under the assumptions that both the nominal closed-loop system and the perturbation L are stable it can be shown from the Nyquist criterion and the identity

$$I + G(I+L) \equiv (I+G)\ [I + (I+G^{-1})^{-1}L] \tag{19}$$

that the perturbed closed-loop system remains stable if

$$\|(I + G^{-1}(j\omega))^{-1}L(j\omega)\| < 1, \qquad \omega > 0 \tag{20}$$

(assuming G^{-1} exists). Again, a weaker condition than (20) but one which directly exposes L is

$$\|L(j\omega)\| < \frac{1}{\|(I + G^{-1}(j\omega))^{-1}\|}, \ \omega > 0 \ . \tag{21}$$

Remark 1: As we noted in Section 1, the above inequalities are tight, i.e., the $<$ cannot be replaced with $\leq$.

Remark 2: Where convenient we shall henceforth drop the "$j\omega$" arguments.

Remark 3: It must be stressed that the results based on

$$\|(I+G^{\pm 1})^{-1}\|\,\|L\| < 1 \qquad (18),\ (21)$$

are weaker than those based on

$$\|(I+G^{\pm 1})^{-1}L\| < 1 \qquad (17),\ (20)$$

since

$$\|(I+G^{\pm 1})^{-1}L\| \leq \|(I+G^{\pm 1})^{-1}\| \cdot \|L\| . \qquad (22)$$

For example, if $L = c(I+G^{\pm 1})$ for some constant c, $|c| < 1$, the differences in the bounds are obvious. In (18), (21) we have

$$\|(I+G^{\pm 1})^{-1}\| \cdot \|L\| = |c| \cdot \kappa(I+G^{\pm 1})$$

while in (17), (20) we have

$$\|(I+G^{\pm 1})^{-1}L\| = |c|$$

and it is possible to have

$$|c| \ll |c| \cdot \kappa(I+G^{\pm 1}) \quad .$$

However, for random perturbations L, (22) is often approximately an equality. To see this, note that a random (dense) L will almost surely be invertible; recall Example 2. It is then easy to show that

$$\frac{\|(I+G^{\pm 1})^{-1}\|}{\|L^{-1}\|} \leq \|(I+G^{\pm 1})^{-1}L\| \leq \|(I+G^{\pm 1})^{-1}\| \cdot \|L\| .$$

Again, since L is random, it will almost surely be well-conditioned (w.r.t. inversion) so that $\|L^{-1}\| \approx \frac{1}{\|L\|}$. Hence,

$$\|(I + G^{\pm 1})^{-1} L\| \approx \|(I + G^{\pm 1})^{-1}\| \cdot \|L\| \ .$$

A related aspect, also worth noting, follows from the inequalities

$$\frac{\|(I + G^{\pm 1})^{-1}\| \cdot \|L\|}{\kappa(I + G^{\pm 1})} \leq \|(I + G^{\pm 1})^{-1} L\| \leq \|(I + G^{\pm 1})^{-1}\| \cdot \|L\| \ .$$

If $(I + G^{\pm 1})$ is reasonably well-conditioned ($\kappa(I + G^{\pm 1})$ near 1) , the majorization (22) will not be a bad overestimate.

Remark 4: By our discussion in Section 1, the appropriate measure of stability robustness is

$$d = \min_{\omega > 0} \frac{1}{\|(I + G^{\pm 1}(j\omega))^{-1}\|} \ . \tag{23}$$

and in the sequel we shall consider methods of efficiently plotting $\frac{1}{\|(I + G^{\pm 1})^{-1}\|}$ as a function of ω. This quantity is familiar from classical sensitivity analysis where it is shown, in the single-input single-output case, that the change in the output of a closed-loop system, due to (additive) perturbations in G (scalar), is reduced by a factor of 1 + G compared with the open-loop effect.

Remark 5: So far we have required nothing of our norm other than $\|I\| = 1$. Of course, a frequently occurring norm in much of the analysis of linear systems is the spectral norm $\|\cdot\|_2$. In that case

$\frac{1}{\|(I+G^{\pm 1})^{-1}\|}$ is the smallest singular value of $(I+G^{\pm 1})$. Let

$$d_q(\omega) = \frac{1}{\|(I+G^{\pm 1}(j\omega))^{-1}\|_q} . \tag{24}$$

We are interested in plotting $d_q(\omega)$ versus ω for large numbers of ω's. We shall see in the sequel that determining $d_2(\omega)$ can be somewhat more expensive to determine than, say $d_1(\omega)$ or $d_\infty(\omega)$. Moreover, note that

$$\frac{1}{\sqrt{m}}\|A\|_2 \leq \|A\|_1 \leq \sqrt{m}\ \|A\|_2 \tag{25}$$

and

$$\frac{1}{\sqrt{m}}\|A\|_2 \leq \|A\|_\infty \leq \sqrt{m}\ \|A\|_2 \tag{26}$$

for $A \in \mathbb{C}^{m\times m}$. Since we are usually most interested in order-of-magnitude estimates of $d_q(\omega)$, $d_2(\omega)$ will lie in a strip sufficiently close to $d_1(\omega)$, for example, to give the same qualitative information. The number m which is the number of inputs/outputs in the system is typically no more than about 10 and is frequently much less.

b. RELATIONSHIPS BETWEEN ADDITIVE AND MULTIPLICATIVE PERTURBATIONS

The following theorem relates additive and multiplicative perturbations. Again, the "$j\omega$'s" will be omitted for convenience and all relations will be assumed to hold for all $\omega > 0$.

Theorem 2: $\left| \|(I+G^{-1})^{-1}\| - \|(I+G)^{-1}\| \right| \leq 1$

Proof: From the identity

$$(I+G^{-1})^{-1} + (I+G)^{-1} \equiv I \tag{27}$$

we have

$$\left| \|(I+G^{-1})^{-1}\| - \|(I+G)^{-1}\| \right| \leq \|(I+G^{-1})^{-1} + (I+G)^{-1}\| = \|I\| = 1.$$

We now get immediately the following useful corollary:

Corollary 1: Assuming that both the nominal closed-loop feedback system of Fig. 1 and the perturbation L are stable then the perturbed system is stable under:

(a) additive perturbations if

$$\|L\| < \frac{1}{1 + \|(I+G^{-1})^{-1}\|} \tag{28}$$

(b) multiplicative perturbations if

$$\|L\| < \frac{1}{1 + \|(I+G)^{-1}\|} \tag{29}$$

Proof: Follows immediately from Theorem 2 noting that

$$\frac{1}{1 + \|(I+G^{\pm 1})^{-1}\|} \leq \frac{1}{\|(I+G^{\mp 1})^{-1}\|} .$$

c. SPECIAL RESULTS FOR THE SPECTRAL NORM

In this subsection we shall present some results related to those in subsections a. and b. but specialized to the $\|\cdot\|_2$ - norm. For

a matrix $H \in \mathbb{C}^{m\times m}$ with singular values $\sigma_1(H) \geq \ldots \geq \sigma_m(H) \geq 0$ we note that $\|H\|_2 = \sigma_1(H)$. If H is nonsingular, $\|H^{-1}\|_2 = \frac{1}{\sigma_m(H)} > 0$. In the $\|\cdot\|_2$ - norm (28) becomes

$$\sigma_1(L) < \frac{\sigma_m(I+G^{-1})}{1+\sigma_m(I+G^{-1})}$$

while (29) becomes

$$\sigma_1(L) < \frac{\sigma_m(I+G)}{1+\sigma_m(I+G)} .$$

We shall make great use in the sequel of the following result of Fan [7].

Theorem 3: Let $A,B \in \mathbb{C}^{n\times n}$. Then

(a) $\sigma_{i+j-1}(A+B) \leq \sigma_i(A) + \sigma_j(B); \quad i \geq 1, \quad j \geq 1$

(b) $\sigma_{i+j-1}(AB) \leq \sigma_i(A)\sigma_j(B); \quad i \geq 1, \quad j \geq 1$

Part (b) of Theorem 3 can be used to relate $\sigma_m(I+G)$ and $\sigma_m(I+G^{-1})$.

Theorem 4: (a) $\frac{1}{\|G^{-1}\|_2}\sigma_m(I+G^{-1}) \leq \sigma_m(I+G) \leq \|G\|_2\sigma_m(I+G^{-1})$

(b) $\frac{1}{\|G\|_2}\sigma_m(I+G) \leq \sigma_m(I+G^{-1}) \leq \|G^{-1}\|_2\,\sigma_m(I+G)$

Proof: Follows immediately from Theorem 3 using

$$I + G^{-1} \equiv G^{-1}(I+G) \quad \text{and} \quad I + G \equiv G(I+G^{-1}) .$$

For the rest of this subsection we shall let H denote either $I + G$ or $I + G^{-1}$ according to whether additive or multiplicative perturbations are appropriate. The next theorem will show how the singular values of $H + L$ can be bounded in terms of $\|L\|_2$ and the singular values of H.

Theorem 5: Suppose $\sigma_k(H) \geq \alpha_k > 0$ for some k, $1 \leq k \leq m$, and $\|L\|_2 \leq \beta$. Suppose further that $\beta < \alpha_k$. Then:

(a) $\sigma_k(I + H^{-1}L) \geq 1 - \frac{\beta}{\alpha_k}$

(b) $\sigma_k(H + L) \geq \alpha_k - \beta$.

(Note: If $k \neq m$, $H + L$ is not necessarily invertible if β is too large.)

Proof: (a) Use $I \equiv I + H^{-1}L - H^{-1}L$ and $A = I + H^{-1}L$, $B = -H^{-1}L$, $i = k$, $j = m-k+1$ in Theorem 3(a) to get

$$\sigma_m(I) \leq \sigma_k(I + H^{-1}L) + \sigma_{m-k+1}(H^{-1}L) .$$

Thus $\sigma_k(I + H^{-1}L) \geq 1 - \sigma_{m-k+1}(H^{-1}L)$

$$\geq 1 - \|L\|_2 \cdot \sigma_{m-k+1}(H^{-1}) \text{ by Theorem 3(b)}$$

$$= 1 - \|L\|_2 \sigma_k(H)$$

$$\geq 1 - \frac{\beta}{\alpha_k} .$$

(b) Use $H \equiv H + L - L$ and $A = H + L$, $B = -L$, $i = k$, $j = 1$ in Theorem 3(a) to get

$$\sigma_k(H+L) \geq \sigma_k(H) - \|L\|_2 \geq \alpha_k - \beta .$$

The case k = m is of special interest in Theorem 5 as it bears directly on our two basic inequalities (18) and (21) of the form

$$\|L\|_2 < \frac{1}{\|H^{-1}\|_2}$$

which are sufficient to guarantee stability of a perturbed closed-loop system. Specifically, if $\|H^{-1}\|_2 \leq \frac{1}{\alpha}$ and $\|L\| \leq \beta$ with $0 \leq \beta < \alpha$, then H + L is invertible and $\|(H+L)^{-1}\| \leq \frac{1}{\alpha - \beta}$ or $\sigma_m(H+L) \geq \alpha - \beta$.

Note that Theorem 5 was expressed in terms of isolating $\|L\|_2$. By analogy with the inequalities (17) and (20) we can also have the following stronger, but perhaps less useful, theorem.

Theorem 6: Suppose $\sigma_{m-k+1}(H^{-1}L) \leq 1 - \delta$ where $0 < \delta < 1$ and $1 \leq k \leq m$. Then:

(a) $\sigma_k(I + H^{-1}L) \geq \delta$

(b) $\sigma_k(H + L) \geq \frac{\delta}{\|H^{-1}\|_2}$

Proof: (a) From the proof of Theorem 5 we have

$$\sigma_k(I + H^{-1}L) \geq 1 - \sigma_{m-k+1}(H^{-1}L) \geq \delta$$

(b) From $I + H^{-1}L \equiv H^{-1}(H+L)$ and Theorem 3(b) we have

$$\sigma_k(I + H^{-1}L) \leq \sigma_k(H + L) \cdot \|H^{-1}\|_2$$

whence $\sigma_k(H+L) \geq \frac{\delta}{\|H^{-1}\|_2}$.

d. SPECIAL RESULTS WHEN $G(s) = C(sI - A)^{-1}B$

In this subsection we shall make use of the fact that the frequency response matrix is of the form

$$G(j\omega) = C(j\omega I - A)^{-1}B$$

Let us further define

$$F(j\omega) = C(j\omega I - A + BC)^{-1}B \tag{30}$$

Recall the Sherman-Morrison-Woodbury formula:

$$(W + XYZ)^{-1} \equiv W^{-1} - W^{-1}X(Y^{-1} + ZW^{-1}X)^{-1}ZW^{-1}$$

assuming the indicated inverses exist. Then it is easy to verify that

$$(I + G(j\omega))^{-1} \equiv I - F(j\omega) \tag{31}$$

and, from (27),

$$(I + G^{-1}(j\omega))^{-1} \equiv F(j\omega) \tag{32}$$

Thus our results in the last section (for example, Theorems 4, 5, and 6) can all be cast in terms of F by noting that

$$\sigma_k(I + G) = \frac{1}{\sigma_{m-k+1}(I - F)} \tag{33}$$

and

$$\sigma_k(I + G^{-1}) = \frac{1}{\sigma_{m-k+1}(F)} \quad . \tag{34}$$

Moreover,

$$\|(I + G)^{-1}\| = \|I - F\| \tag{35}$$

and

$$\|(I + G^{-1})^{-1}\| = \|F\| \tag{36}$$

for any of the norms we have been considering (in particular, $k = m$ in (33) and (34)). Use of (31) and (32) results in an apparent savings in the number of linear systems to be solved (i.e., number of inversions) and we shall exploit this fact in the next section.

4. COMPUTATIONAL PROBLEMS

a. COMPUTATION OF FREQUENCY RESPONSE MATRICES

As we have seen above, an object of considerable interest in studying the robustness of stability of linear systems is a graph of $\frac{1}{\|(I + G^{\pm 1}(j\omega))^{-1}\|}$ as a function of ω. When $G(j\omega) = C(j\omega I - A)^{-1}B$ we saw that $\|(I + G(j\omega))^{-1}\| = \|I - F(j\omega)\|$ and $\|(I + G^{-1}(j\omega))^{-1}\| = \|F(j\omega)\|$ where $F(j\omega) = C(j\omega I - A + BC)^{-1}B$. Thus, regardless of the norm used, a quantity of the form

$$C(j\omega I - H)^{-1}B \tag{37}$$

must first be computed. We shall assume throughout this and the next subsection that: (i) $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{m\times n}$, $H \in \mathbb{R}^{n\times n}$ are given

(ii) $n \geq m$

(iii) (37) is to be evaluated for a large number, N, of values of ω; typically $N >> n$.

Rather than concentrate on exact operation counts, which may be fairly meaningless anyway, we shall give only order-of-magnitude estimates. It will be seen that the bulk of the computational load rests on evaluating matrices of the form (37) and so we shall focus initially on that problem.

If $A \in \mathbb{R}^{n\times n}$ is dense, the most efficient evaluation of $C(j\omega I - A)^{-1}B$ by an LU factorization of A, solution of m triangular systems to get $(j\omega I - A)^{-1}B$, and finally a matrix multiplication, requires approximately $\frac{1}{3}n^3 + \frac{1}{2}mn^2 + m^2n$ multiplications (and a like number of additions; we shall henceforth count only multiplications). This figure, when multiplied by N, represents a rather large amount of computation.

If A is initially transformed, however, the computational burden can be reduced quite considerably. If T is a similarity transformation on A we have

$$C(j\omega I - A)^{-1}B \equiv CT(j\omega I - T^{-1}AT)^{-1}T^{-1}B \quad .$$

Let us define

$$H = T^{-1}AT$$

and agree, for convenience to still label CT, $T^{-1}B$ the transformed C and B matrices, respectively, as C, B respectively. We now have the problem of evaluating

$$C(j\omega I - H)^{-1}B$$

where H may now be in such a form that $(j\omega I - H)^{-1}$ can be computed in less than $O(n^3)$ operations. For example, A can always be reduced to upper Hessenberg form by (stabilized) elementary transformations ($\frac{5}{6}n^3$ multiplications) or by orthogonal transformations ($\frac{5}{3}n^3$ multiplications). These transformations are very stable numerically and, while $O(n^3)$, are performed only once at the beginning of the calculations. The resulting linear system to be solved - for N different values of ω — now has an upper Hessenberg coefficient matrix and can be solved in approximately $\frac{1}{2}mn^2$ multiplications. Moreover, Hessenberg systems can be solved very accurately with the growth factor in Gaussian elimination bounded above by n; see [8]. Computing $C(j\omega I - H)^{-1}B$ still requires an additional m^2n multiplications. Neglecting the initial transformation and determination of CT and $T^{-1}B$, the Hessenberg method requires approximately

$\frac{1}{2}mn^2 + m^2n$ multiplications (for each value of ω), a considerable savings over the $O(n^3)$ algorithm if $n >> m$.

Of course, other transformations T are possible. One possibility is to reduce A to upper triangular (Schur) form by means of orthogonal similarities. This is considerably more expensive than reduction to upper Hessenberg but, again, need only be done once at the beginning. However, the resulting linear system to be solved at each step is upper triangular and so still requires $O(mn^2)$ multiplications. Because of potential difficulties with multiple eigenvalues of A there seems to be little real advantage gained by this procedure. Substantial savings could be gained though if the eigenstructure of A were such that it was diagonalizable by a reliably computable T. Since this involves consideration of the essentially open numerical problems associated with computing invariant subspaces we shall not pursue the details here. But assuming such a transformation were possible, $C(j\omega I - D)^{-1}B$ with D diagonal, could be computed with approximately $mn + m^2n$ multiplications for each value of ω. Attractive as this appears, the potential for severe ill-conditioning of the eigenproblem associated with A render this latter method unreliable as a general-purpose approach. We shall subsequently consider only the Hessenberg method.

The analysis above has been done under the assumption that complex arithmetic was performed. We now outline how $G = C(j\omega I - H)^{-1}B$ might be determined using only real arithmetic. The matrix H is assumed to be in upper Hessenberg form. We wish to solve first

$$(j\omega I - H)Z = B \tag{38}$$

Then

$$G = CZ .$$

Suppose $Z = X + jY$ where $X, Y \in \mathbb{R}^{n\times m}$. Upon equating real and imaginary parts in (38) we get the following order 2n real system to determine X and Y:

$$\begin{pmatrix} -H & -\omega I \\ \omega I & -H \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} B \\ 0 \end{pmatrix} \tag{39}$$

Thus $X = \frac{1}{\omega} HY$ and $Y = -\omega(\omega^2 I + H^2)^{-1}B$. The matrix $(\omega^2 I + H^2)$ will be invertible if $(j\omega I - H)$ is invertible. Note that $(\omega^2 I + H^2)$ is no longer upper Hessenberg but is almost in the sense of having two rather than one nonzero subdiagonal. Its shape is wholly typified for n = 5 by the matrix

$$\begin{pmatrix} x & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \\ 0 & x & x & x & x \\ 0 & 0 & x & x & x \end{pmatrix} .$$

Linear systems involving matrices of this type can be solved using approximately n^2 multiplications. We summarize the Hessenberg method using real arithmetic:

(i) Reduce A to upper Hessenberg form H, transform B and C, and compute H^2; this step is done only once.

(ii) Solve $(\omega^2 I + A^2)Y = -\omega B$ for Y.

(iii) Compute $X = \frac{1}{\omega} HY$.

(iv) Compute $G = (CX) + j(CY)$.

Step (ii) requires approximately mn^2 multiplications, step (iii) requires approximately $\frac{1}{2} mn^2$, and step (iii) approximately $m^2 n$. The total number of multiplications is approximately $\frac{3}{2} mn^2 + m^2 n$.

Storage requirements for the Hessenberg method with real arithmetic are approximately double those for complex arithmetic.

b. COMPUTATION OF ROBUSTNESS MEASURES

We have seen above that quantities of the form (37) can be reliably evaluated in $O(mn^2)$ operations. There then remains the problem of determining (35) or (36).

Case 1: $\| \cdot \|_2$

For (35), the singular value decomposition (SVD) of $I + G(j\omega)$ can be computed for each value of ω. Each SVD typically requires approximately $6m^3$ multiplications. The smallest singular value is then the quantity of interest. For (36), inversion of G can be avoided by finding the SVD of $F(j\omega)$, again in approximately $6m^3$ multiplications. The inverse of the largest singular value of F is then the quantity of interest.

Case 2: $\| \cdot \|_1$ or $\| \cdot \|_\infty$

Use of either of these norms in (35) or (36) involves negligible computation as compared to Case 1, namely about m^2 additions and absolute

values and m-1 arithmetic comparisons.

In both cases, the additional work required is usually small compared with $O(mn^2)$ especially if $n >> m$. However, if m is large relative to n, significant savings can be realized in using $\|\cdot\|_1$ or $\|\cdot\|_\infty$ rather than $\|\cdot\|_2$. In fact, using our previous approximate operation counts for the Hessenberg method and setting $n = km$, we have

$$\rho = \frac{\text{work per value of } \omega \text{ using } \|\cdot\|_2}{\text{work per value of } \omega \text{ using } \|\cdot\|_1 \text{ or } \|\cdot\|_\infty} \simeq \frac{k^2 + 2k + 12}{k^2 + 2k}$$

Note though that $\rho \approx \frac{k^2 + 2k + 24}{k^2 + 2k}$ if singular directions are also computed.

In the event A (or A - BC) can be successfully diagonalized as mentioned in Section 4.a. the potential savings in avoiding $\|\cdot\|_2$ are somewhat greater. In fact, we then have

$$\rho \approx \frac{k+6}{k}$$

(or $\rho \approx \frac{k+12}{k}$ if singular directions are also computed).

The above comparisons are only approximate and should in no way be construed as definitive statements. The purpose of this section is to merely introduce certain aspects of the numerical computations and suggest further avenues of exploration. A great deal of numerical experimentation remains to be done. Reliable software such as LINPACK [9] for linear systems will be of great benefit in this research.

5. CONCLUSIONS

We began this paper with an attempt at a "formal" definition of robustness. We then applied the definition to the problem of robustness of stability of linear systems as discussed in [1]. The cases of both additive and multiplicative perturbations were discussed and a number of relationships between the two cases were given. Finally, a number of computational aspects of the theory were discussed including a proposed new method for evaluating general transfer or frequency response matrices. The new method is numerically stable and efficient, requiring only $O(mn^2)$ operations to update for new values of the frequency parameter rather than $O(n^3)$

A number of interesting research areas suggest themselves in this work. One such area is that of constrained perturbations. For example, in our basic problem we were concerned with the nearness to singularity of a nonsingular matrix $A \in \mathbb{C}^{nxn}$. If the admissible perturbations E are somehow constrained for one reason or another, for example E upper triangular, the usual bound on $\|E\|$ for which A + E is singular but E is "dense" may be overly pessimistic. Related to this is the fact that our bounds were derived for the "worst case". The size of perturbations allowed in a linear system to ensure continued closed-loop stability may very well be larger than we have derived if inputs to the system are constrained in certain directions.

We have concentrated in this paper on the analysis of linear control systems. There are many interesting — and difficult — synthesis problems, however. For example, can A, B, C be chosen to assign certain singular values of $I + G^{\pm 1}$? What is the effect of changes in B or C on the

behavior of $I + G^{\pm 1}$? Can a matrix K be determined so that $I + (GK)^{\pm 1}$ has certain singular values?

On the computational side, more research needs to be done on updating parametric problems. That is, suppose we have a matrix (say, $G(j\omega)$) which depends "in a rank m way" on a parameter ω. When ω changes how can various quantities be updated efficiently?

Finally, as mentioned in Section 4.b., a great deal of numerical experimentation is necessary to get a qualitative feel for the numbers in determining robustness measures.

Acknowledgment: It is a pleasure to acknowledge several stimulating conversations with Gunter Stein and Nils R. Sandell, Jr. which motivated much of the research reported in this paper.

6. REFERENCES

[1] Stein, G., Robust Stability of Linear Systems — Engineering Motivation, these Proceedings.

[2] Wonham, W.M., Linear Multivariable Control. A Geometric Approach, Springer-Verlag, Berlin, 1974.

[3] Kahan, W., Numerical Linear Algebra, Can. Math. Bull., 9(1966), 757-801.

[4] Willems, J.C., The Analysis of Feedback Systems, M.I.T. Press, Cambridge, MA, 1971.

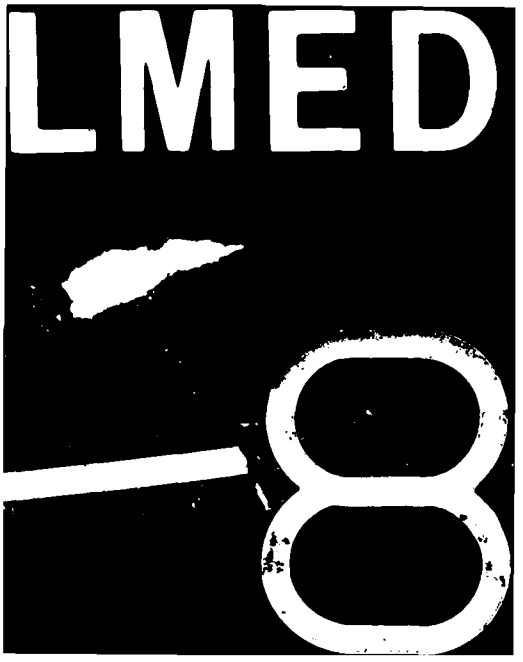
[5] Desoer, C.A., Perturbation in the I/O Map of a Non-Linear Feedback System Caused by Large Plant Perturbation, J. Franklin Institute, 306(1978), 225-235.

[6] Rheinboldt, W.C., On Measures of Ill-Conditioning for Nonlinear Equations, Math. Comp., 30(1976), 104-111.

[7] Fan, K., Maximum Properties and Inequalities for the Eigenvalues of Completely Continuous Operators, Proc. Nat. Acad. Sci., 37(1951), 760-766.

[8] Wilkinson, J.H., Error Analysis of Direct Methods of Matrix Inversion, J. Assoc. Comput. Mach., 8(1961), 281-330.

[9] Dongarra, J.J., C.B. Moler, J.R. Bunch, and G.W. Stewart, LINPACK User's Guide, SIAM, Philadelphia, 1979.


LMED
8